

칼라와 에지 정보를 이용한 영상 검색

심 성 오, 최 태 선

광주과학기술원 기전공학과 신호 및 영상 처리 실험실

전화 : 062-970-2419 / 핸드폰 : 016-325-7926

Content Based Image Retrieval using Color and Edge

Seong-O Shim, Tae-Sun Choi

Dept. of Mechatronics, K-JIST

E-mail: seongo@kjist.ac.kr

요약

영상 검색을 위한 기존의 칼라 히스토그램 방법은 영상의 형상 정보를 포함하고 있지 않다. 본 논문에서는 로컬 형상 정보인 에지 정보에 칼라 정보를 접목시켰다. 각각의 세 종류의 에지(수평 에지, 수직 에지, 비방향성 에지) 주위의 픽셀들에 대한 칼라 분포를 구한 후 그 각각의 칼라 분포를 기초로 구한 두 영상간의 거리 정보를 이용하여 영상들간의 유사도를 구했다. 따라서 본 논문에서 제안한 유사도는 한 종류의 에지 타입(비 방향성 에지)을 이용한 방법에 비해 매칭 오류를 줄일 수 있었다. 실험 결과, 제안된 방법이 기존의 칼라 히스토그램과 에지 히스토그램을 이용한 방법에 비해 향상된 성능을 보였다.

I. 서론

Content-Based Image Retrieval (CBIR)의 목적은 오직 영상 정보만을 이용하여 사용자가 제시한 query

영상과 유사한 영상을 검색하는 것이다. 따라서 CBIR에서는 효과적이면서도 강인한 영상 기술자를 정의하는 것이 아주 중요하다. 칼라는 가장 우세한 시각 구별자로 간주가 되며, 영상 내 전역 칼라 분포를 나타내는 칼라 히스토그램이 CBIR에서 많이 사용되고 있다[1, 2]. 칼라 히스토그램은 피사 물체의 이동과 회전에 강인한 반면 어떠한 공간 정보도 포함하고 있지 않다. 전혀 다른 영상이 비슷한 칼라 분포를 가질 수 있으며, 영상 내 물체의 큰 외관상의 변화는 쉽게 히스토그램을 변화시킬 수 있다. 따라서 CBIR에서는 영상 내 형상정보가 필수적이며, 이 형상정보를 칼라 정보와 결합시키는 방법이 많이 연구되고 있다[3, 4, 5, 6, 7]. 이론적으로 형상정보를 추출하는 최적의 방법은 영상 내 물체를 분할하는 것이다. 하지만 영상 분할 과정은 아주 복잡하고 시간이 많이 걸리며 영상 내 특정 물체가 없다면 엉뚱한 결과가 나올 수 있다.

본 논문에서는 영상분할 과정을 거치지 않고, 영상의 로컬 형상 정보인 에지를 칼라와 접목시키는 방법을 시도하였다. 각각의 세 종류의 에지(수평 에지, 수직 에지, 비 방향성 에지) 주위의 픽셀들에 대한 칼라 분포

를 구한 후 그 각각의 칼라 분포를 기초로 구한 두 영상간의 거리 정보를 이용하여 영상들간의 유사도를 제안하였다.

II. 에지 칼라 히스토그램

에지는 로컬 영상 정보로써 영상 내 일반적인 영상 정보를 포함하고 있다[8]. 에지는 영상 인지에 있어 중요한 역할을 하기 때문에 영상검색에서 영상 기술자로써 많이 사용되고 있다. 에지 히스토그램이 그 예로써, 다섯 종류의 에지(네 개의 방향성 에지와, 한 개의 비방향성 에지) 들의 분포를 나타낸다[9]. 하지만 에지 히스토그램은 회전에 대한 강인성이 없다. 영상의 회전은 각각의 에지 구성요소 들을 이동시킬 뿐 아니라, 각 에지의 멤버쉽도 변화시킨다[12].

2.1 에지 주위에서의 칼라 분포

영상 내 칼라의 변화는 칼라 에지에서 일어난다. 따라서 칼라 에지 주위 픽셀들에 대한 칼라 분포는 영상 전체에 대한 칼라 분포와 비슷하다.

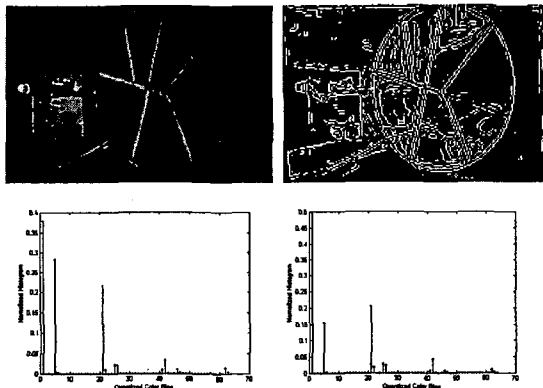


그림 1. 영상 전체에 대한 칼라분포와 에지 주위에 대한 칼라 분포의 비교

그림 1에서 보듯이 두 히스토그램은 값에 있어서의 변화를 제외하고는 비슷한 경향을 보이고 있다. 이는 영상 내 큰 영역을 차지하는 칼라는 에지 주위의 픽셀들만을 고려 시에도 큰 영역을 차지하기 때문이다.

2.2 에지 칼라 히스토그램

본 논문에서는 두 방향성 에지(수평 에지, 수직 에지)와 한 개의 비 방향성 에지(수평 에지와 수직 에지의 결합)주위 픽셀들에 대해 칼라 히스토그램을 구하였다. 따라서 제안된 영상 기술자는 2차원인 (n, m) 형태를 취하며, 이때 n 은 세 종류의 에지를 그리고 m 은 n 번째 에지의 칼라 분포의 양자화된 칼라 수를 나타낸다. 앞으로 이것을 에지 칼라 히스토그램으로 부르겠다.

에지 칼라 히스토그램을 구하는 첫번째 과정은 영상 내 에지를 구하는 것이다. 우선 R, G, B 각 채널별로 두 Sobel 연산자(수평 연산자와 수직 연산자)를 적용하여 임계 값을 취한 뒤 OR 연산을 하여 두 개의 이진 에지 영상을 구한다. 임계 값은 각 채널별로 최대값의 15%로 하였다. 그리고 구한 두 에지 영상을 더하여 비방향성 에지 영상을 구한다. 에지 주위 픽셀을 취하기 위해 각 세 에지 영상을 한번 dilate 시킨다. 마지막으로 원 영상으로부터 앞에서 구한 세 에지 영상에서 이진 값 1을 갖는 픽셀들에 대해서 총 세 개의 칼라 분포를 구한다.

E_j 를 j 번째 에지에 속하는 픽셀들의 집합이라 하고 n_j 를 다음과 같이 정의한다.

$$n_j = \left| \{ p \mid p \in E_j \} \right| \quad (1)$$

여기서 $|\cdot|$ 은 원소의 수를 나타낸다. 영상 I 에서 j 번째 에지에 대한 칼라 히스토그램은,

$$h_j(j, k) = n_j \cdot \text{Pr} \left[p \in I_k \mid p \in E_j \right] \quad (2)$$

로 정의한다. 여기서 k 는 양자화된 칼라의 인덱스를 나타내며, 에지를 나타내는 j 는 수평, 수직, 비 방향성 에지 순으로 정렬한다. 총 m 개의 칼라로 양자화했다면 $k \in [m]$ 로 나타낼 수 있다. 마지막으로 scaling invariant를 위해 정규화를 하여 에지 칼라 히스토그램을 다음과 같이 정의한다.

$H_I(j, k) = h_i(j, k) / n_j = \Pr_{p \in I} [p \in I_i | p \in E_j]$ (3)
 $j \in [3]$ 이기 때문에 히스토그램 $H_I(\cdot)$ 의 차원은 $3 \times m$ 이다.

III. 유사도 측정

본 논문에서는, 사용자가 제시한 query 영상 Q 와 영상 데이터 베이스로부터의 영상 I 사이의 유사도 $S(Q, I)$ 를 에지 칼라 히스토그램 $H(\cdot, \cdot)$ 에 기초하여 정의 하였다. 우선 거리 벡터 $D(Q, I) = (d_1, d_2, d_3)$ 를 계산하였다. 여기서 d_j 는 다음과 같이 정의된다.

$$d_j = \|H_Q(j, k) - H_I(j, k)\| \quad (4)$$

여기서 $\|\cdot\|$ 는 *Euclidean distance*를 나타낸다. 마지막으로 유사도 $S(Q, I)$ 는 다음과 같이 정의하였다.

$$S(Q, I) = \frac{w_1 d_1 + w_2 d_2 + w_3 d_3}{w_1 + w_2 + w_3} \quad (5)$$

여기서 w_j 는 각 d_j 에 부여된 가중치다. 비슷한 영상은 낮은 유사도 값을 가지며, 동일한 영상은 유사도가 0이 된다.

비슷한 영상끼리의 d 값은 낮은 값을 갖지만, 다른 영상간의 d 값도 낮은 값을 갖을 수 있다(매칭 오류). 에지 종류별로 거리 d 를 계산하여 식 (5)에 제시한 유사도를 적용함으로써 이러한 매칭 오류를 개선할 수 있다. 왜냐하면, 다른 영상이 모든 세 개의 d 값이 낮은 확률은 낮기 때문이다.

본 논문에서는 거리 벡터의 세 구성요소 들에 똑같은 가중치를 부여하였다($w_1 = w_2 = w_3 = 1$). 실험 결과 d_3 를 이용한 매칭 결과가 d_1, d_2 를 이용한 매칭 결과에 비해 조금 우수한 성능을 보였다. 따라서 d_3 에 약간 높은 가중치를 부여할 수 있겠다.

IV. 성능 측정

$\{Q_1, \dots, Q_q\}$ 를 query 영상들의 집합이라 하고, Q_i 을 query Q_i 에 대한 올바른 응답이라고 정의 하자. 제안된 방법의 성능을 평가하기 위해 [5]에서 제시한 방법을 이용하였다.

1. r -measure: 모든 query 영상에 대한 올바른 응답의 평균 rank. 즉, $r = (1/q) \sum_{i=1}^q \text{rank}(Q_i')$.

2. p_1 -measure: recall 1에 대한 평균 precision. 즉, $p_1 = (1/q) \sum_{i=1}^q 1/\text{rank}(Q_i')$.

3. *recall vs. scope*: R_1, R_2, \dots, R_t 를 query 영상 Q 에 대해 검색된 상위 t 개의 영상이라 하고, 그 각각은 영상 Q 와의 관련성에 따라 labeling한다. *recall* r 은 *scope* $s (> 0)$ 에 대해 다음과 같이 정의된다.

$$r(s) = |\{R_i | \text{relevant}(R_i) = \text{true}, 1 \leq i \leq s\}|$$

낮은 r -measure 값과 높은 p_1 -measure 값이 좋은 성능을 나타낸다.

V. 실험 결과

에지 칼라 히스토그램을 이용한 제안된 방법은 총 6,000개의 영상으로 이루어진 데이터 베이스를 이용하여 실험되었다. 이 데이터 베이스는 80개의 query 영상과 그에 상응하는 응답들을 포함한다. Query 영상들은 영상의 변위, 회전, 형태 변화, 그리고 빛의 강도의 변화 등 다양한 상황을 나타내도록 선택되었다.

이전 방법들과 비교하기 위해, 기존의 칼라 히스토그램을 사용한 유사도(S_{C_hist}), 에지 히스토그램을 사용한 유사도(S_{E_hist}), 그리고 제안한 칼라 에지 히스토그램을 사용한 유사도(S_{EC_hist})에 대해 ranking을 계산하였다. S_{C_hist} 와 S_{E_hist} 는 히스토그램간의 Euclidean distance를 이용하여 계산하였다. 앞의 식 (5)에서 볼 수 있듯이 S_{EC_hist} 는 d_1, d_2 , 그리고 d_3 에 대한 함수이다 ($S_{EC_hist} = f(d_1, d_2, d_3)$).

이전의 방법들과 제안한 알고리즘의 성능은 Table

1에 비교되었다.

	r -measure	p_1 -measure
S_{E_hist}	99.8	2.48
S_{C_hist}	85.8	2.04
d_1	31.5	3.34
d_2	67.4	2.69
d_3	29.4	3.35
S_{EC_hist}	22.6	3.86

Table 1. 제안된 방법과 이전 방법들 사이의 성능 비교

Table 1에 나타난 바와 같이, 제안된 에지 칼라 히스토그램 방법은 기존의 칼라 히스토그램을 이용한 방법, 그리고 에지 히스토그램을 이용한 방법들에 비해 우수한 성능을 보였다. 여기서 주목할 만한 점은 d_1 , d_2 , 그리고 d_3 를 이용한 ranking도 기존의 방법들에 비해 우수한 성능을 나타냈다는 점이다. d_1 , d_2 , 그리고 d_3 중에서는 어느 하나도 다른 하나보다 월등히 나은 성능을 보이지는 않았고, 단지 d_3 를 이용한 ranking이 다른 두 개를 이용한 ranking에 비해 조금 나은 결과를 나타냈다. 그리고 제안된 유사도 S_{EC_hist} 를 이용한 ranking이 예상했던 대로 d_1 , d_2 , 그리고 d_3 를 이용한 ranking보다 우수한 성능을 나타냈다.

그림 2에 다양한 scope에 대한 relevant 영상의 개수를 비교하였다.

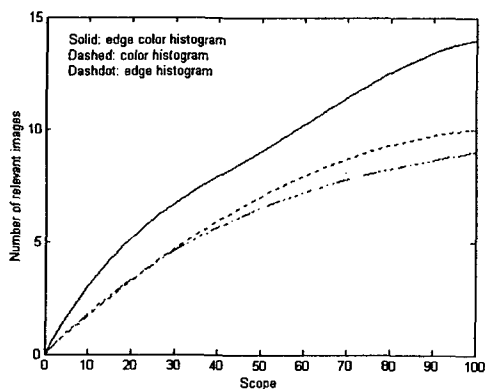


그림 2. 다양한 방법들에 대한 recall vs. scope

VI. 결론

본 논문에서는 에지 주위의 픽셀들에 대한 칼라 분포를 나타내는 에지 칼라 히스토그램을 이용한 영상 검색을 시도 하였다. 제안된 에지 칼라 히스토그램은 칼라와 로컬 형상 정보를 포함하고 있기 때문에 기존의 칼라 히스토그램과 에지 히스토그램 방법에 비해 향상된 검색 결과를 보였다. 또한 에지를 종류별로 구분하여 계산된 거리들을 이용한 유사도는 매칭 오류를 줄임으로써 검색 결과의 성능 향상을 가져왔다.

참고 자료

- [1] M. Swain and D. Ballard, "Color indexing", *International Journal of Computer Vision*, 7(1):11-32, 1991.
- [2] M. Flicker *et al.*, "Query by image and video content: The QBIC system", *IEEE Computer*, 28(9):23-32, 1995.
- [3] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval", *IEEE Workshop on Applications of Computer Vision*, pp. 96-102, 1996.
- [4] M. Stricker and A. Dimai, "Color indexing with weak spatial constraints", *Proc. SPIE Storage Retrieval Still Image Video Databases IV*, vol. 2670, 1996, pp. 29-40.
- [5] J. Huang, S. R. Kumar, M. Mitra, W. Zhu, and R. Zabih, "Image indexing using color correlograms", *Proc. Conf. Computer Vision Pattern Recognition*, 1997, pp. 762-768.
- [6] Ying Deng, B. S. Manjunath, Charles Kenney, Michael S. Moore, Hyundoo Shin, "An efficient color representation for image retrieval", *IEEE Transactions on Image Processing*, vol. 10, Jan. 2001, pp. 140-147.
- [7] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain, "Content-Based Image Retrieval at the end of the Early Years", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, No. 12, Dec. 2000, pp. 1349-1380.
- [8] Sami Brandt, Jorma Laaksonen and Erkki Oja, "Statistical Shape Features in Content-Based Image Retrieval", *Proceedings of 15th International Conference on Pattern Recognition*, vol. 2, 2000, pp. 1062-1065
- [9] International Organization for Standardization ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Audio, ISO/IEC JTC1/SC29/WG11 N4031, Singapore, March 2001.
- [10] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, John Wiley & Sons, 1987.
- [11] James Hafner *et al.*, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, No. 7, July. 1995, pp. 729-736.
- [12] A. K. Jain, A. Vailaya, "Image retrieval using color and shape", *Pattern Recognition*, 29(8), pp. 1233-1244, 1996